Running Head: GAZE AVERSION AND TRANSIENT KNOWLEDGE

Gaze Aversion During Children's Transient Knowledge and Learning

Gwyneth Doherty-Sneddon, Fiona G. Phelps, and Lesley Calderwood

University of Stirling

Correspondence should be addressed to G. Doherty-Sneddon, Department of Psychology, University of Stirling, FK9 4LA, U.K. Email: gds1@stir.ac.uk

Abstract

Looking away from an interlocutor's face during demanding cognitive activity can help adults and children answer challenging mental arithmetic and verbal-reasoning questions (Glenberg, Schroeder, & Robertson, 1998; Phelps, Doherty-Sneddon, & Warnock, 2006). While such "gaze aversion" (GA) is used far less by 5-year old school children, its use increases dramatically during the first years of primary education, reaching adult levels by 8-years of age (Doherty-Sneddon, Bruce, Bonner, Longbotham, & Doyle, 2002). Furthermore GA increases with increasing mental demands, with high levels signalling that an individual finds material being discussed challenging but remains engaged with it (Doherty-Sneddon et al., 2002; Doherty-Sneddon & Phelps, 2006). In the current study we investigate whether patterns of gaze and gaze aversion during children's explanations can predict when they are in states of transient knowledge (Karmiloff-Smith 1992; Goldin-Meadow, Kim, & Singer, 1999). In Study 1, fifty-nine 6-year-olds took part and completed a "Time Task" along with periodic teaching intervention to improve their comprehension of telling the time. Some children improved immediately, whereas others did so more gradually. The gradual improvers showed the highest levels of GA, particularly when they were at an intermediate level of performance. In Study 2, thirty-three 6-year-old children completed a balance beam task (Pine & Messer, 2000). Children who improved the representational level of their explanations (Karmiloff-Smith, 1992) of this task with training used more GA than those who did not. Practical implications for teaching and for recognizing transient knowledge states are discussed.

Gaze Aversion During Children's Transient Knowledge and Learning

During difficult cognitive activity (e.g. remembering information, thinking of an answer to a question, speech-planning, speaking) we often close our eyes, look up at the sky, or look away from the person we are in conversation with (Glenberg et al., 1998; Doherty-Sneddon et al., 2002). Several studies report ways in which adults "switch off" from environmental stimulation (both live faces and other sorts of visual displays) by averting their gaze, in order to concentrate on cognitive tasks (e.g. Beattie, 1981; Day, 1964; Glenberg et al., 1998)—the cognitive load explanation of gaze aversion (GA). Previc, Declerck, and Brabander (2005) propose that such gaze shifts during thinking reflect the particular mental activity involved in a task and that this GA is related to the difficulty of active thought processes. Indeed, adults' use of GA has been shown to increase in response to increasingly effortful cognitive activity (e.g. Glenberg et al., 1998). When questioned face to face, 8-year-old children have also been shown to look away more when answering difficult questions compared with easy ones (Doherty-Sneddon & Phelps, 2005; Doherty-Sneddon et al., 2002). GA is therefore a potentially useful cue during pedagogical interactions, since it gives a non-verbal indication of a child's level of understanding and concentration.

Additionally, it has been shown that increasing the amount that people avert their gaze improves question-answering performance (Glenberg et al., 1998; Phelps, Doherty-Sneddon, & Warnock, 2006), suggesting that it serves an important role in problem solving. For example, Phelps, et al. (2006) trained 5-year-old children to increase the proportion of time they spent averting their gaze during questioning and found that the accuracy of responses to questions increased significantly as a consequence of this training. In addition, Doherty-Sneddon, Bonner, and Bruce (2001) found that children's performance on visuospatial memory tasks was improved by encouraging them to look away from an experimenter's face. So, GA may be a useful indicator that attention has been shifted from external sources of information to allow internal reflection upon a question's answer, with the problem's level of difficulty measurable by the proportion of time spent averting gaze by a participant. This shift to internal processing, in turn, facilitates and has a functional benefit for the problem solving in hand (Phelps et al., 2006).

The theoretical framework that we propose is that visual communication cues are informative and hence carry a processing cost that can interfere with task-relevant processing resources. Our program of GA research started with the finding that young children did more poorly on tasks involving a high load on visuospatial working memory when they looked at faces concurrently (Doherty-Sneddon et al., 2000; Doherty-Sneddon et al., 2001). Given that GA occurs in response to cognitive difficulty of a range of different types of tasks (not just those loading heavily on visuospatial working memory), including autobiographical memory questions, mental arithmetic, and verbal reasoning questions (Doherty-Sneddon & Phelps, 2005), it is likely that other aspects of working memory are involved. The face-to-face interference effects we have found, the benefits of GA on task performance (Phelps et al., 2006), and recent pilot work showing exceptionally high proportions of time spent averting gaze even on easy tasks for children with ADHD (Doherty-Sneddon, Phelps, & Calderwood, 2007) all support the view that GA serves to facilitate working memory function. In the current study we expand on our previous work, which has shown GA to be an external indicator of concentration and mental effort. We investigate whether particular patterns of GA can be used to judge not only mental effort, but when children are in states of transient knowledge.

Over the last two decades, one of the most important contributions to our understanding of how human mental development proceeds is Karmiloff-Smith's Representational Redescription (RR) model (Karmiloff-Smith, 1992). She proposes that much of cognitive development involves the recoding (or redescription) of what is initially implicitly held knowledge into that which is increasingly consciously appreciated (explicit) and ultimately open to verbalization. Movement from one stage to the next involves transition in knowledge structures. For example, during implicit understanding, children accomplish tasks with little conscious understanding of how they do so.

With development, children begin to abstract out the possible rules surrounding their success and begin to apply a theory to their achievement. In the case of language learning, Karmiloff-Smith (Karmiloff-Smith & Inhelder, 1974) proposes that children over-generalize a variety of rules e g. applying the “ed” rule in past tense forms of irregular verbs (goed instead of went; swimmed instead of swam). Pine and Messer (1999) developed this model further and through empirical observations proposed six stages of recoding of information (there were four in the original), as documented in Table 1.

INSERT TABLE 1 ABOUT HERE

It could be argued that solving a problem with explicitly held knowledge is more cognitively demanding than implicit solving (via involvement of both working memory, particularly central executive resources, and long term memory rules to finding a problem’s solution). Our earlier work has linked GA with high cognitive load, and we therefore investigate whether we find evidence of increased GA (associated with high cognitive load) when children evidence a move through the stages of recoding of information between implicit and explicit knowledge. In Study 1 we investigate whether children who deepen their understanding on a balance beam problem (irrespective of external task performance) use more GA than children who do not evidence a change in knowledge state.

Related to this account of transitional knowledge as a cognitively demanding time is the proposal that transitional knowledge states can be identified via unstable competence on a task. Alibali and Goldin-Meadow (1993) report that some children gain competence on a task directly, going from zero competence to complete competence; these children they refer to as “skippers.” In contrast, other children (the majority) acquire competence in a more gradual, stepwise manner, going from zero competence, to an unstable level of competence to full task competence. Goldin-

Meadow, Nusbaum, Garber, and Church (1993) propose that transitional knowledge states (associated with an unstable level of competence) are especially cognitively demanding, requiring more processing resources than when children are competent or incompetent on a task. We therefore predict that when children experience intermediate states in their performance level, they will use more GA than when they are in the stable states and more overall than the immediate improvers. In Study 2 we look at whether transitional task performance is associated with high levels of GA.

There are, therefore, different accounts of what transitional knowledge is or how knowledge levels change. We predict that knowledge development is cognitively demanding, involving high levels of individual processing resources. The point of this paper is not to support or refute a particular theoretical account of how knowledge changes, but to provide evidence that when it does change, this is a cognitively demanding time. We hypothesize that while children evidence either an increase in explicitness of underlying knowledge (Study 1) or a gradual increase in task performance (Study 2), the proportion of time spent in GA will be especially high.

In Study 1, a group of 6-year-olds engaged in a learning task, the balance beam task (Pine & Messer, 2003), to establish whether the pattern of increased GA during task improvement (found in earlier work) could be replicated. The interesting feature of this task is that it allows a differentiation between task "success" per se and task comprehension or learning. So, for example, when comprehension of the task is low, children are often able implicitly but correctly to balance symmetrical beams, but unable to explain how they did it. As comprehension progresses, performance drops initially, although their explanations and reasoning about their strategies improve. See Table 1 for definitions of stages of progress on the beam task and the stages of mental representation redescription with which these are associated (Karmiloff-Smith 1992).

Study 1: Balance Beam Task

Method

Participants

Participants were thirty-three 6-year-olds, 16 girls and 17 boys, with a mean age of 6 years 3 months (age range = 5 years 8 months – 6 years 10 months). Children were recruited through local schools and took part only upon the acquisition of written parental permission. Eight of the children failed to complete all stages of the experiment, leaving 25 children in the final sample.

Design

A pre-test, training, and post-test design was used. The dependent variables were: the amount of pre-test to post-test improvement in ability to balance symmetric and asymmetric beams; levels of explanations given by the children of how the beams balanced or did not balance, which were coded using the 6-level coding scheme developed by Pine and Messer (2003), ranging from implicit to explicit comprehension; and the percentage of time spent averting gaze while explaining why beams did or did not balance in the pre-test and post-test.

Procedure

Balance task. There were three symmetrical beams with the same number of blocks on either side (one with no blocks on either side, one had one block on either side, and one had two blocks on each side). These beams balanced in the geometric center. A more difficult version of the task involves asymmetric beams (Pine & Messer, 2000). There were three asymmetrical beams in the current study with a different number of blocks on each end (one with two blocks at one end and no blocks at the other end, one with three blocks at one end and no blocks at the other end, and one with three blocks at one end and one block at the other end). All of these beams balanced off-center.

Protocol. Children were tested individually in a quiet area of the school. The child sat across from the experimenter at a table where the fulcrum and the beams were placed. A video camera was placed directly behind the experimenter so that the child's head and shoulders could be recorded and gaze aversion could be coded later. The children engaged in a pre-test, a training phase, and a post-test.

The child was first informed that he or she was going to be talking about balance, and was asked if he or she could explain what balancing meant. Then the child was shown the fulcrum and told that he or she would be given some beams to try to balance on the fulcrum. The child was told to try to make each beam stay level, so that it did not tip off to one side. For the pre-test, children were asked to balance the six beams (three symmetric; three asymmetric) one by one, and after the attempt to balance each beam, were asked to explain why it did or did not balance. The experimenter gave no feedback at this stage. Any child who could balance all six beams correctly and gave full explicit explanations of how the beams balanced in terms of weight and length did not proceed any further. This was the case for only one child.

Training. After the pre-test, children completed a training phase where the experimenter modelled how to balance one of the symmetrical and one of the asymmetrical beams and encouraged the child to think about how they balanced. No specific information was provided at this stage about the compensatory function of length and weight for the asymmetrical beams or about the need to have the same amount of weight on each side of the beam so that it would balance. Children observed the experimenter balancing the beams and then were asked to talk about how they had balanced. For example after watching the experimenter balance a symmetrical beam, they were asked questions like, “How did I get that one to balance? Why did I put it in the middle? What have we got on this side, what have we got on this side?” to try to get them to think about the same weights being on each side. After an asymmetrical beam, they were asked “How did I get that one to balance? Why doesn't it go in the middle? What have we got on this side, what have we got on this side?” The children then had another go at balancing all the beams, and then the experimenter again showed them how to balance one symmetrical and one asymmetrical beam, but this time explicitly talked about weight and length, e.g., “This one balances in the middle because we have the same size of weights on each side and the same length of beam on each side." The asymmetrical beams don't balance in the middle because we have a heavier weight on this side, and if we put it in the middle it will fall down to this side, so we need to make the other side longer so that we have the same weight on each side. So, how do we get the same weight on each side? Well, what have we got on this side? (child normally responds 'blocks') and what have we got on this side? (child normally responds 'blocks'). But what else have we got on this side?" The experimenter would then point to the length of the beam and explain that this also gives weight. The experimenter explained that for symmetrical beams the lengths had to be the same on both sides, but that for asymmetrical beams, the side with fewer blocks would have to be longer so that you would have the same weight overall on each side. This modelling was done as naturally as possible with each child, although specific contributions from the experimenter came from a predetermined standardized script.

Pine and Messer (2000) found significant improvements on the post-test of the beam task when children observed an adult modelling the correct solution and were encouraged to explain how the experimenter balanced the beams. Children in this condition did better on the post-test than children who only observed the experimenter and did not discuss how the beams were balancing. Pine and Messer also report different levels of expertise associated with different forms of explanation given by children. These are defined in Table 1.

Children then carried out the first post-test, where they were again given the six beams to balance and were again asked to explain how/why the beams did or did not balance. Testing finished here for any child who successfully balanced all six beams and gave full explicit explanations in terms of length and weight. For those with only partial success in performance and/or explanations, training was repeated and a second post-test given.

Classification of children's progress. In order to gain a clearer impression of how GA changes as children's knowledge on the balance task develops, children were divided into two groups, according to the level at which their explanations were coded at the pre-test and post-test stages. Any child who showed consistent improvement as the experiment progressed (such as being classified as implicit at the pre-test and then classified as abstraction-verbal at the first post-test and as explicit at the second post-test) was classified as an improver. Children who did not show any improvement in their explanations as the experiment progressed (e.g., continued to give implicit explanations at each stage) or who regressed (e.g., those who were classified as abstraction-verbal at the pre-test and then gave implicit explanations at the post-test) were classified as non-improvers (only one non-improver actually regressed). Table 2 gives data on criteria as well as levels of explanations given by both groups at pre-test and post-test. This shows that while the starting point on levels of explanation given by all children were similar, about two thirds of them showed rapid increases in the sophistication of their explanations, while the others remained very static.

INSERT TABLE 2 ABOUT HERE

Gaze aversion coding. Thinking time was defined as the period of time when the experimenter finished asking the question and the child began speaking a response. The GA dependant measure was the percentage of thinking time spent averting gaze. An independent coder analyzed the video records of 10% of the children randomly sampled for GA. The coders measured thinking time as well as GA time in seconds and from this ascertained the percentage of thinking time scores for each trial coded. There was a very strong correlation between the GA data point scores produced by each of the two coders, $r(71) = .90$, $p < .001$, evidencing very good agreement on the GA coding.

Results

Task Performance

The mean number of symmetric and asymmetric beams successfully balanced at the pre-test, post-test-1, and post-test-2 is displayed in Table 3 (maximum = 3 for both kinds of beam). A three-way ANOVA was carried out: task difficulty (2 levels: symmetric beam, asymmetric beam) and training stage (3 levels: pre-training; post-training-1; post-training-2) within participant variables, group (2 levels: improvers; non-improvers) was a between participant variable. The ANOVA revealed a significant effect of training on task performance, F (2,46) = 48.85, p < .001, η_p^2 = .68, with better performance after training than before (mean pre-training = 1.69 correct responses; mean post-training-1 = 2.60 correct responses; mean post-training-2 = 2.85 correct responses). Task difficulty also had a significant effect on performance, F (1,23) = 64.32, p < .001, η_p^2 = .74, with better performance on the symmetric beams (2.95 correct responses) than on the asymmetric beams (1.80 correct responses). There was no significant main effect of group, F (1,23) = 3.24, p = .09, η_p^2 = .12, (mean improver performance = 2.51 correct responses; mean non-improver performance = 2.24). There was a significant interaction between training and task difficulty, F (2,46) = 42.38, p <.001, η_p^2 = .65. Means are given in Table 3. Simple effects analyses showed that training stage had a significant effect for asymmetric beams, F(2,48) = 54.34, p < .001, η_p^2 = .70, but not for symmetric ones F(2,48) = 2.87, p = .07, η_p^2 = .11). Post-hoc t-tests showed that performance was better on symmetric beams compared with asymmetric ones at every training stage {pre-test: t(24) = 9.53, p < .001; post-test-1: t(24) = 3.48, p < .01; post-test-2: t(24) = 2.30, p < .05}. Finally, there was a significant interaction between task difficulty and group, F (1,23) = 4.24, p = .05, η_p^2 = .16. Post-hoc t-test revealed a significant group effect, with improvers better than non-improvers on the asymmetric beams, t(23) = 2.00, p < .05, but not on the symmetric beams, t(23) = .38, p > .05.

INSERT TABLE 3 ABOUT HERE

In summary, improvers did better than non-improvers on the harder version of the task, although both groups significantly increased their performance when the experimenter modelled the correct technique.

Gaze Aversion

The proportion of time children averted their gaze while explaining why the beams did or did not balance was calculated from the videos and is presented in Table 4 for each stage of the experiment (pre-test, post-test-1 and post-test-2).

INSERT TABLE 4 ABOUT HERE

A three-way ANOVA was used to analyse the percentage of time children spent averting gaze during the explanations. The ANOVA included task difficulty (2 levels: symmetric beam, asymmetric beam) and training stage (3 levels: pre-training; post-training-1; post-training-2) as within participant variables. Group (2 levels: improvers; non-improvers) was a between-participant variable. Task difficulty had a significant impact on the percentage of time children spent averting their gaze, with more GA occurring as children explained asymmetric beams $F(1,23) = 4.59$, $p <.05$, $\eta_p^2 = .17$ (mean GA on symmetric beams = 55.4% of explanation time; mean on asymmetric beams = 59.1%). Group also had a significant effect on GA, $F(1,23) = 4.85$, $p < .05$, $\eta_p^2 = .17$ (mean GA for improvers was 64.23%; for non-improvers, 50.29%). There was a significant interaction between task difficulty and group, $F(1, 23) = 5.98$, $p < .05$, $\eta_p^2 = .21$. Means are given in Table 4. Post-hoc t-tests showed that only the non-improvers increased GA with increasing task difficulty, $t(26) = 4.06$, $p <.001$. In addition, improvers used significantly more GA than non-improvers at both levels of task difficulty (on symmetrical beams: $t(73) = 3.90$, $p < .001$; on asymmetrical beams: $t(73) = 2.59$, $p < .05$).

Discussion

Study 1 replicates a finding from a number of earlier studies (e.g. Doherty-Sneddon et al., 2002; 2004; 2005; 2007)—that GA increases as task demands get harder. This was only the case for the non-improvers, however, as improvers' levels of GA were already high, even on the easier version of the task. Furthermore, although non-improvers increased their GA with increased task difficulty, their GA was still significantly lower than the improvers' GA during the harder version of the task. The current work shows that children who improve their level of task comprehension (defined by their explanations of their task solutions) use more GA than those who do not. Of course this data represents a snapshot of children's behaviour when they have just improved their task comprehension. While out of the scope of the current study, it may be that with increased expertise/practice on the task, proportion of time spent averting gaze will fall as task demands reduce. So, while explicit processing is likely to be more demanding than implicit processing because of the greater attentional resources required of it, practice and expertise will also play a role in determining absolute level of task demand.

Study 2: Time Task

Method

Participants

Participants were fifty-nine 6-year-olds. Seven of the children, who either (a) answered all the questions correctly at all stages of difficulty, and so there was no training required, or (b) failed to make any progress at all after training (going from 0 correct to 0 correct) were removed from the analysis. In the remaining 52, there were 18 girls and 34 boys, mean age 6 years 5 months (age range, 5 years 9 months – 7 years 1 month). Children were recruited through local schools and took part only upon the acquisition of written parental permission.

Design

A pre-test, training, and post-test design was used. The dependent variables were the amount of pre-test to post-test improvement in ability to tell the time and the percentage of time spent averting gaze while thinking of a response in the pre-test and post-test.

Procedure

Children were tested individually in a quiet area of the school. The child sat across from the experimenter at a table. A video camera was placed directly behind the experimenter so that the child's head and shoulders could be recorded and gaze aversion could be coded later. Children were first informed that they were going to be talking about time and that they would have to work out what time it was, depending on where the big and little hands of a clock were pointing. Four levels of difficulty were used—o'clock (level 1), half past (level 2), quarter past (level 3) and quarter to (level 4). The task and the different levels of difficulty were proposed by the class teachers, as telling time was a topic they were currently working on and that the children had not yet mastered. There were three pre-test questions at each level. All children started at level 1 and were asked the following questions: If the big hand is at 12 and the little hand is at 2 (4, 9), what time would that be? If the child answered all three of these questions correctly, he or she progressed to the next level and so on, until reaching a level where at least one of the three questions was answered incorrectly. A child reaching this stage received training from the experimenter. The training involved viewing pictures of clocks with different times that represented the appropriate level of difficulty (e.g., 1 o'clock, 4 o'clock and 7 o'clock if they were trained at level 1). The experimenter pointed out where the big and little hands were on each picture, emphasizing where the big hand was in each picture and explaining how this determined the time for that particular level (e.g., "When the big hand is at 12, that always means it is something o'clock."). The child was then given a clock and was asked to set the hands to different times, according to the level that they had reached (e.g., "Can you put the hands to 6 o'clock?"). Feedback was given until the child set the time correctly three times in a row without help. Then the child was asked three post-test questions (e.g., "If the big hand is at 12 and the little hand is at 7 (1, 11) what time would that be?"). If any of the post-test-1 questions were answered incorrectly, then the training phase was repeated and the post-test phase (post-test-2) was repeated.

The proportion of time children averted their gaze while thinking about responses to the test questions was calculated at the level where they required training. At this level their GA was measured at pre-test, post-test-1, and post-test-2. Thinking time was defined as the period of time when the experimenter finished asking the question and the child began speaking a response. The GA dependant measure was the percentage of thinking time spent averting gaze. It should be noted that task materials were put out of sight during questioning. An independent coder analysed 10% of the video records for GA (as in Study 1). There was a significant correlation between the two coders' GA data points, $r(26) = .79$, $p <.001$, indicating a high level of agreement.

Results

Gradual Improvers

Task performance. Accuracy of responses at pre-test, post-test-1, and post-test-2 was calculated as a percentage of correct responses. Twenty-eight of the children went from 0% accuracy on pre-test to 100% accuracy immediately on post-test-1. These children were classed as immediate improvers (in the Alibali and Goldin-Meadow (1993) terminology they would be called skippers). The remaining 23 children exhibited a much more gradual increase in task competence, requiring the additional second training session before achieving full competence at post-test-2. These children were classed as gradual improvers. Because of the different levels of testing sessions (only two for the immediate improvers and three for the gradual improvers), the data for the two groups of children are analyzed separately.

A one-way ANOVA with repeated measures on session (3 levels: pre-test, post-test-1, post-test-2) was used to analyze the accuracy scores. The ANOVA revealed a significant change in accuracy across session, $F(2,44) = 131.26$, $p < .001$, $\eta_p^2 = .86$ {mean pre-test = 4.35% (*SD* 15.25) of responses were accurate; at post-test-1, 33.33% (*SD* 28.00); post-test-2, 94.20% (*SD* 12.92)}. Post-hoc t-tests showed that each session differed significantly from the others: pre-test/post-test-1, $t(22) = 4.60$, $p < .001$; post-test-1/post-test-2, $t(22) = 10.02$, $p < .001$; pre-test/post-test-2, $t(22) = 20.36$, $p < .001$. These children therefore went from very low performance before instruction to achieving about a third of responses accurately after one session of instruction to almost complete task competence after their second round of instruction.

Gaze aversion. A one-way ANOVA was used to analyze the percentage of thinking time these children spent averting their gaze in each of the testing sessions. The ANOVA revealed a significant effect of testing session, $F(2,44) = 3.63$, $p < .05$, $\eta_p^2 = .14$ {mean pre-test = 70.55 (*SD* 20.83); post-test-1 = 82.30 (*SD* 16.36); post-test-2 = 70.00 (*SD* 26.83)}. Post-hoc t-tests showed that GA increased significantly between pre-test and post-test-1, $t(22) = 2.78$, $p < .05$, but then fell again between post-test-1 and post-test-2, $t(22) = 2.53$, $p < .05$. There was no significant difference between GA at pre-test and post-test-2 ($t(22) = .09$, $p = ns$).[1]

Taken along with the accuracy data, we see that for these gradual improvers, GA does, indeed, increase while children are experiencing a change in task expertise compared with tasks that are entirely out of reach of their abilities or where they have achieved task competency. We argue that this reflects the cognitive effort of having mental strategies to deal with the problem in hand that are not yet easily applied.

Immediate improvers

Task performance. With children who showed immediate improvement on the task, the picture is very different. Their accuracy increased immediately from no to complete task competence (0% to 100% across only two sessions). In addition, their GA did not change across sessions, $t(27) = .15$, $p > .05$ (mean pre-test = 73.89 (*SD* 20.66); mean post-test-1 = 73.23 (*SD* 16.52) and was similar to the level of GA exhibited by the gradual improvers at pre-test and post-test-2 sessions.[2]

Gaze aversion. When children are either at no or complete competence, their GA is similar (around 73% of thinking time). In contrast, when children experience partial competence (evidenced by a gradual improvement in performance), their GA peaks at around 82% of thinking time, compared with approximately 70% of the time when they cannot do the task at all or have already mastered it.

Discussion

These results suggest that especially high proportions of GA are associated with transient knowledge states. Clearly the children are averting their gaze for the majority of their thinking time (as we have shown in previous research). Whether teachers would be able to detect a 12% increase in GA in natural classroom interactions remains to be seen (such a change translates to a 420ms increase in a 3.5 second episode). What is important here is that the relative increase in GA indicates the cognitive difficulty of transitional knowledge states. During such times children are likely to benefit most from a "freedom" to avoid visual contact with another person.

General Discussion

It was hypothesized that when children experience a transient knowledge state (as defined by either an unstable level of performance plus a gradual increase in performance on task or an increased ability to explicitly reason about a task), their GA will be especially high. The data support this hypothesis. In Study 1 children whose quality of explanations improved (by showing an increased level of explicit understanding of the balance beam) used more GA than those whose explanations did not show a growing understanding. Interestingly, on this task both improvers and non-improvers balanced more beams correctly following modelling of the correct technique. The children who were beginning to understand and to *explain* the technique (improvers) used the greatest proportion of GA. We propose that this reflects the greater processing resources (perhaps in relation to working memory) involved on the task by children whose understanding is on the increase. In Study 2 around half of the children were immediate improvers on the time concept. The others exhibited a more gradual improvement with instruction, and it was these children who used the highest proportion of GA. In addition, these children increased their GA significantly during their transient or intermediate performance level, compared with when their performance was at floor or when it was at ceiling. Whether teachers would be able to detect this 12% increase in GA in natural classroom interactions remains to be seen (which translates to a 420ms increase in a 3.5 second episode in the currently used Time Task). The important point is that children whose performance (Study 2) or comprehension (Study 1) is improving avert their gaze the most. These children may tolerate face-to-face contact the least and may indeed suffer the most from culturally imposed notions that lead adults to say things like, "Look at me when I'm speaking to you."

This pattern of results makes sense, given the account of transient knowledge provided by Goldin-Meadow (1999). She demonstrates that transient knowledge states are especially demanding by showing the lack of resources available for a secondary, unrelated task attempted when children's knowledge is undergoing change. The present study shows evidence of high cognitive demand in the transient state by demonstrating that high proportions of time spent averting gaze are associated with it. Of course, transient states are not defined solely by performance, that is, by partial task competence. In the beam task, children were able to copy an adult model and achieve greater task performance. It is only when their understanding of the task increases and they are truly engaged in solving the task themselves that the mental load increases and hence the GA. This finding relates to our earlier claims that a relatively high proportion of time spent in GA reflects task engagement (e.g. Phelps et al., 2006).

High GA during transitional knowledge states may reflect heavy loading on working memory resources, particularly executive components. The face-to-face interference effects we have found (Doherty-Sneddon et al., 2000; Doherty-Sneddon et al., 2001), the benefits of GA on task performance (Phelps et al., 2006) and recent pilot work showing exceptionally high amounts of GA even on easy tasks for children with ADHD (Doherty-Sneddon et al., 2007), all support the view that GA serves to facilitate working memory function. An important area of research to advance is investigating the relationship between working memory function, patterns of GA, and transitional knowledge.

An important question is whether the current results will translate to knowledge acquisition within foundational domains. We have used mathematics questions in our gaze aversion paradigm in many studies and have established solidly that gaze aversion increases with cognitive difficulty of mathematics questions. In addition, we have found this pattern of GA in response to cognitive load in a number of different types of task, including biographical memory and verbal reasoning. In Study 1 of the current paper we show that this pattern of increasing GA during difficult activity is also found in a very different task (balancing beams). Furthermore, time reading is part of the Scottish national curriculum and was suggested for the study by the participants' teachers. We are therefore confident that the pattern of results we have found in the time and balance beam tasks will translate to other domains, including those of foundational knowledge, and that children experiencing transitional knowledge within any domain will exhibit a peak in GA.

Even experienced teachers are little better than novices at detecting whether or not children have understood something on the basis of reading non-verbal cues (Jecker, Maccoby, & Breitrose, 1965. Teachers' recognition of these strategies is crucial to effective scaffolding of learning (Wood, Bruner, & Ross, 1976). If GA can be used reliably to judge children's readiness to learn, this will have significant implications for teacher training. We have recently shown that while teachers interpret GA correctly in some contexts (e.g., they associate it with answering hard questions compared with easy), they often misinterpret it, for example associating it with a child having given up thinking (Phelps et al., 2006). Furthermore, and very importantly, teachers do not seem to use GA to judge wait time (the time between asking a question and waiting for the child's response), and instead frequently interrupt the child's thinking, which is very detrimental to learning (Doherty-Sneddon & Phelps, 2007). The current work shows that high proportions of time spent on GA are associated with progress in learning. In contrast, low levels of GA are likely to reflect less comprehension and progress.

Future research work should extend what we know about GA as cognitive load management in typical development to populations of children with special learning needs. We are currently piloting work with children and young people with autism, ADHD, and Williams syndrome. Patterns of results indicate that GA promises to provide new and important insights into the cognitive as well as social functioning of these groups. For example, our pilot work with children with ADHD suggests that these children do not tolerate face gaze (i.e., they avert their gazes almost all the time), even during easy tasks. In the pilot study the children were engaged and performed well and in line with matched controls. What they could not do was maintain face contact while thinking. It may be that the particular deficits in working memory capacity associated with ADHD (Brocki, Randall, Bohln & Kerns, 2008), can explain this “face-to-face shut-down.” This is beyond the scope of the current paper but is a line of research our team is currently pursuing. Understanding the adaptive and constructive nature of GA is of utmost importance for teaching typical but especially atypically developing children, where a “look at me when I’m talking to you” approach may well be counter-productive.

Author’s Note

Gwyneth Doherty-Sneddon is senior lecturer and director of research in the Department of Psychology, University of Stirling. Fiona Phelps is now research associate in the School of Psychology, Cardiff University. Lesley Calderwood is an associate of the Department of Psychology, Glasgow Caledonian University.

References

Alibali, M.W. & Goldin-Meadow, S. (1993). Gesture-speech mismatch and mechanisms of learning: What the hands reveal about a child's state of mind. *Cognitive Psychology*, *25*, 468-523.

Beattie, G. W. (1981). A further investigation of the cognitive interference hypothesis of gaze patterns during conversation. *British Journal of Social Psychology*, *20*, 243-248.

Brocki, K.C., Randall, K.D., Bohln, G., Kerns, K.A. (2008). Working memory in school-aged children with attention-deficit/hyperactivity disorder combined type: Are deficits modality specific and are they independent of impaired inhibitory control? *Journal of Clinical and Experimental Neuropsychology*, *30*,749-759.

Day, M. E (1964). An eye movement phenomenon relating to attention, thought and anxiety. *Perceptual and Motor Skills, 19,* 443-446.

Doherty-Sneddon, G., McAuley, S., Bruce, V., Langton, S., Blokland, A., & Anderson, A. (2000). Visual signals and children's communication: Negative effects on task outcome. *British Journal of Developmental Psychology*, *18*, 595-608.

Doherty-Sneddon, G., Bonner, L., and Bruce, V. (2001). Cognitive demands of face monitoring: Evidence of visuospatial overload. *Memory and Cognition*, *29*, 909-919.

Doherty-Sneddon, G., Bruce, V. Bonner, L., Longbotham, S. & Doyle, C. (2002). Development of gaze aversion as disengagement from visual information. *Developmental Psychology*, *38*, 438-445.

Doherty-Sneddon, G. (2004). Don't look now I'm trying to think: Children's eye gaze and cues to comprehension. *The Psychologist*, *17*, 82.

Doherty-Sneddon, G. & Phelps, F. (2005). Gaze aversion: A solution to cognitive or social difficulty? *Memory and Cognition*, *33*, 727-733.

Doherty-Sneddon, G. Phelps, F., & Clarke, J. (2007). Development of gaze aversion: Qualitative changes over the early school years. *British Journal of Developmental Psychology*, 25, 513-526.

Doherty-Sneddon, G. & Phelps, F. (2007). Teachers' responses to children's gaze aversion. *Journal of Educational Psychology*, *27*, 91-107.

Doherty-Sneddon, G., Phelps, F., & Calderwood, L (2007, August). Gaze aversion in ADHD. Paper presented at the meeting of British Psychological Society: Developmental Section Plymouth, UK.

Glenberg, , A. M., Schroeder, J. L, & Robertson, D. A. (1998). Averting the gaze disengages the environment and facilitates remembering. *Memory and Cognition*, *26*, 651-658.

Goldin-Meadow, S., Kim, S., & Singer, M. (1999).What the teacher's hands tell the student's mind about math. *Journal of Educational Psychology*, *91*, 720-730.

Goldin-Meadow, S., Nusbaum, H., Garber, P. & Church, R. B. (1993). Transitions in learning: Evidence for simultaneously activated strategies. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 92-107.

Jecker, J.D., Maccoby, N., & Breitrose, H.S. (1965). Improving accuracy in interpreting non-verbal cues of comprehension. *Psychology in the Schools*, *2*, 239-244.

Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental perspective on cognitive science*. Cambridge, MA: MIT Press.

Karmiloff-Smith, A. & Inhelder, A. (1974). If you want to get ahead get a theory. *Cognition*, *3*, 195–212.

Phelps, F., Doherty-Sneddon, G., & Warnock (2006). Functional benefits of children's gaze aversion during questioning. *British Journal of Developmental Psychology*, 24, 577-588.

Pine, K., J., & Messer, D. (1999). What children do and what children know: Looking beyond success using Karmiloff-Smith's R framework. *New Ideas in Psychology*, *17*, 17-30.

Pine, K., J., & Messer, D. (2000). The effects of explaining another's actions on children's implicit theories of balance. *Cognition and Instruction*, *18*, 37-54.

Pine, K., J., & Messer, D. (2003). The development of representations as children learn about balancing. *British Journal of Developmental Psychology*, *21*, 285-301.

Previc, F.H., Declerck, C., and de Brabander, B., (2005). Why your "head is in the clouds" during thinking: The relationship between cognition and upper space. *Acta Psychologica*, *118*, 7-24.

Wood, D., Bruner, J., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, *17*, 89-100.

Footnotes

1. The raw amount of time it took children on average to perform the task was 3.48 seconds. There was no significant difference in the time taken between pre- and post-test sessions although there was a trend for the children to get faster after training, $F(2,46) = 3.24$, $p = .09$ (mean pre-test: 4.06 seconds; post-test-1: 3.48 seconds; post-test-2: 2.91 seconds).

2. The immediate improvers got significantly faster at the task between pre-test and post-test, $t(27) = 3.06$, $p < .001$ (mean pre-test: 4.48 seconds; mean post-test: 2.40 seconds).

Table 1

Criteria for Representational Levels on the Balance Beam Task (from Pine & Messer, 1999)

Implicit level	The child can successfully balance at least 2 of both symmetric and asymmetric beams. However no consistent strategy shown and no ability to verbalise understanding or strategy e.g. says “I just did it”.
Implicit transition	Child is able to balance no more than 1 of each beam type. A centre strategy used with all beams placed on the fulcrum around their midpoint. Explanations similar to implicit level.
Abstraction non-verbal	Success with symmetric beams but fails on all, or all but one, asymmetric. Centre strategy as at implicit transition. Child does not explain their centre strategy and may state that asymmetric beams cannot be balanced.
Abstraction verbal	Success with symmetric beams but not asymmetric. Centre strategy as at abstraction non-verbal but can now explain the centre strategy for symmetric beams.
Explicit transition	Success on both types of beam. Explains strategies for both types of beam: a centre strategy for symmetric and e.g. “It has to go a bit over to the side” for asymmetric. However

	no explanation of weight and distance.
E3 (explicit 3)	Success on both types of beam. Can explain functional relationship between weight and distance.

Table 2

Classification of Children as Improvers or Non-improvers

	Classification	Primary pre-test explanation type	Secondary explanation type at pre-test	Post-te
Improvers	Consistent improvement between pre-test and post-test e.g. implicit->abstraction-verbal	14/16 were at abstract verbal level 2/16 were at explicit transition level	3/14 combined with E3 level; 4/14 combined with explicit transition; 2/14 combined with implicit level	14/16 v level o
Non-improvers	No improvement between pre-test and post-test e.g. implicit->implicit level explanations	8/9 were at abstract verbal level 1/9 implicit level	3/9 with implicit level; 1 explicit transition; 1 explicit combination	7/9 stil 1/9 exp 1/9 im

Table 3

The Number of Symmetric and Asymmetric Beams Correctly Balanced at Each Stage for Each Group (with Standard Deviations)

	Pre-test		Post-Test-1		Post-Test-2	
	Improver	Non-improver	Improver	Non-improver	Improver	Non-improver
Type of Beam						
Symmetric	2.81 (0.54)	2.89 (.33)	3.00 (0.00)	3.00(.00)	3.00 (0.00)	3.00 (.00)
Asymmetric	0.94 (1.29)	0.11 (.33)	2.38 (0.96)	2.00 (1.32)	2.94 (0.25)	2.44 (.73)

Table 4

The Proportion of Time Spent Averting Gaze when Explaining Why the Beams Did or Did Not Balance

	Pre-test	Post-Test-1	Post-Test-2	Overall
Symmetric:				
Improvers	62.23 (17.25)	63.74 (16.84)	67.51 (15.74)	64.50 (4.25)
Non-improvers	43.82 (22.18)	47.15 (26.17)	48.03 (25.59)	46.33 (5.66)
Asymmetric:				
Improvers	62.74 (10.83)	63.90 (12.78)	65.26 (16.12)	63.97 (3.60)
Non-improvers	53.65 (17.48)	55.19 (21.87)	53.91 (18.29)	54.25 (4.80)